REMARKS

Applicant appreciates the Examiner's thorough consideration provided the present application. Claims 1-7 and 9-16 are now present in the application. Claims 1 and 9 have been amended and claims 8 and 17 have been cancelled by this Amendment. Claims 1 and 9 are independent. Reconsideration of this application, as amended, is respectfully requested.

Claim Rejections

Claims 1-4, 8-13 and 17 stand rejected under 35 U.S.C. § 102(b) as being anticipated by D'Hoore (U.S. Patent No. 6,085,160); claims 5 and 13 stand rejected under 35 U.S.C. § 103(a) as being unpatentable over D'Hoore in view of Burns (U.S. Patent No. 5,454,106); claims 6, 7, 15 and 16 stand rejected under 35 U.S.C. § 103(a) as being unpatentable over D'Hoore in view of Waibel ("Interactive Translation of Conversational Speech", IEEE 1996).

Applicant respectfully traverses the rejections of claims 1-17 of the present application. Before addressing the details of specific rejections, Applicant notes that there are fundamental differences between the disclosure of D'Hoore and that of the present invention.

The present invention generally relates to a system and method for multi-lingual speech recognition, comprising: a speech modeling engine, receiving and transferring a mixed multi-lingual speech signal into a plurality of speech features; a speech search engine, coupled to the speech modeling engine, receiving the speech features, and locating and comparing a plurality of candidate data sets corresponding to the speech features, <u>referring the connecting sequences of the speech features and a speech rule database</u>, to find match probability of a plurality of candidate speech models of the candidate data sets; and a decision reaction engine, coupled to the speech search engine, selecting a plurality of resulting speech models corresponding to the speech features according to the match probability from the candidate speech models to generates a speech command.

Relatively, D'Hoore discloses a language independent speech recognition system comprising: a speech pre-processor which receives input speech and produces a speech-related signal representative of the input speech; a database of acoustic hidden Markov models which represent subword units in each of a plurality of languages, wherein any subword unit that is common to two or more of the plurality of languages is represented by a single common acoustic hidden Markov model; a language model which characterizes a vocabulary of recognizable words and a set of grammar rules; and a speech recognizer which compares the speech-related signal to the acoustic hidden Markov models and the language model, and recognizes the input speech as a specific word sequence of at least one word.

The Examiner has taken the position that the connecting sequences of the speech features and a speech rule database have been disclosed in D'Hoore since the biphone acoustic models are referred to the connecting sequences while the language model are referred to the speech rule database. The acoustic model is made up of the patterns of speech sounds, such as phonemes (the smallest units of sound) or words. The language model generally incorporates the set of words and sentences that are allowed and expected within the context of the application.

Here, the connecting sequences may follow some specific connection rules in a particular application, such as an ID or address, and the acoustic model is made up of the patterns of speech sounds that do not relate any connecting sequences following connection rules. Additionally, the language model generally incorporates the set of words and sentences, which is not related to speech rules, so the language model should not be corresponded to the speech rule database.

Although D'Hoore has disclosed most technical features of the present invention, the locating and comparison of the candidate data sets further referring the connecting sequences of the speech features and the speech rule database is not disclosed.

As described, the disclosure of the citation is obviously different than that of the present invention. Thus, the present invention is novel based on D'Hore and should be allowable.

Rejection of Claims 1 and 9

As described, the amended claims 1 and 9 focus on the locating and comparison of the candidate data sets by further referring to the connecting sequences of the speech features and the speech rule database, limitations of which are not disclosed in D'Hoore.

Thus, claims 1 and 9 are novel based on the features of D'Hoore and should be allowable.

Rejection of Claims 2 and 10

Referring to claims 2 and 10, the present invention teaches that the speech models are characterized by diphone models.

D'Hoore teaches "phoneme-like subword units such as biphones and triphones based on Hidden Markov Models (HMMs)".

The present invention and D'Hoore are implemented by different models, respectively and, further, Applicant believes that the technical features of claim 2 are distinguishable over D'Hoore. Thus, the limitations of the present invention are not disclosed in the D'Hoore citation. Therefore, claims 2 and 10 are novel based on the features of D'Hoore and should be allowable.

Rejection of Claims 4 and 12

Referring to claims 4 and 12, the present invention teaches that the multi-lingual model database comprises multi-lingual context-speech mapping data.

D'hoore teaches "context dependent acoustic models are trained and used for recognition".

The multi-lingual context-speech mapping data has no relationship with the context dependent acoustic models and generation steps thereof are different that the context dependent

acoustic models, which is generated by a multi-lingual baseform generation engine and a cross-lingual diphone model generation engine.

Thus, the limitations of the present invention are not disclosed in D'Hoore and claims 4 and 12 are novel based on the features of D'Hoore and should be allowable.

Rejection of Claim 14

Referring to claim 14, the present invention teaches that selection and combination further comprises the steps of: fixing left contexts of the multi-lingual baseforms and mapping right contexts of the multi-lingual baseforms to obtain a mapping result; fixing right context and mapping the left contexts of the multi-lingual baseforms to obtain the mapping result if the right contexts of the multi-lingual baseforms mapping fails; and obtaining the multi-lingual context-speech mapping data according to the mapping result.

D'Hoore teaches "The training procedures for single language and multi-language acoustic models both use standard training techniques ... The training process begins by training context independent models using Viterbi training of discrete density HMMs ... Based on the class information, context dependent phoneme models are constructed. Next, the context dependent models are trained using Viterbi training of discrete density HMMs. The context dependent and context independent phoneme models are merged, and then, lastly, badly trained context dependent models are smoothed with the context independent models. Such acoustic model training methods are well-known within the art of speech recognition".

Although D'Hoore discloses single language and multi-language acoustic models, training context independent models using Viterbi training of discrete density HMMs, merging the context dependent and context independent phoneme models, and smoothing the trained context dependent models with the context independent models, but those steps are not related to that of the present invention, comprising fixing left contexts and mapping right contexts to obtain a mapping result, fixing right context and mapping the left contexts to obtain the mapping

PCL/jcg

result, and obtaining the multi-lingual context-speech mapping data according to the mapping result.

Thus, the limitations of the present invention are not disclosed in D'Hoore citation and claim 14 is novel based on the features of D'Hoore and should be allowable.

Rejection of Claims 5 and 13

Referring to claims 5 and 13, the present invention teaches a multi-lingual baseform mapping engine, comparing a plurality of multi-lingual query commands to obtain a plurality of multi-lingual baseforms, and a cross-lingual diphone model generation engine, selecting and combining the multi-lingual baseforms to generate the multi-lingual context-speech mapping data.

D'Hoore teaches creating the acoustic models, constructing the acoustic model of a particular phoneme based on speech from multiple languages , using phoneme-like subword units such as biphones and triphones based on Hidden Markov Models (HMMs), training context independent models using Viterbi training of discrete density HMMs, constructing context dependent phoneme models, training the context dependent models using Viterbi training of discrete density HMMs, merging the context dependent and context independent phoneme models, and smoothing trained context dependent models. The described process has no relationship with comparing the multi-lingual query commands and selecting and combining the multi-lingual baseforms operations of the present invention. Therefore, Applicant believes the assertion by the Examiner that the technical features of the present invention have been disclosed in D'Hoore to be unreasonable.

Burns teaches receiving an input query which can be in the form of text input, converting the user input query into a system recognizable format, and retrieving information from a database using natural language (NL) queries and graphical interfaces and displays. Also, the described process has no relationship with comparing a plurality of multi-lingual query

commands to obtain a plurality of multi-lingual baseforms. The only common aspect is using a query command, which cannot represent that the comparing step has been disclosed in Burns.

Thus, the limitations of the present invention are not disclosed in D'Hoore and Burns and claims 5 and 13 are inventive based on the features of D'Hoore and Burns and should be allowable.

Rejection of Claims 6 and 15

Referring to claims 6 and 15, the present invention teaches that the multi-lingual model database comprises a plurality of multi-lingual anti-models.

D'Hoore teaches using standard training techniques, training context independent models using Viterbi training of discrete density HMMs, constructing context dependent phoneme models, training the context dependent models using Viterbi training of discrete density HMMs, merging the context dependent and context independent phoneme models, and smoothing the trained context dependent models with the context independent models. Additionally, the Waibel citation teaches applying the garbage models to model non-stationary noises.

Although D'Hoore and Waibel apply various models, the present invention applies the multi-lingual anti-models different from those model in D'Hoore and Waibel. Different models may result in different effects for speech recognition. Therefore, Applicant believes the assertion by the Examiner that the models of the present invention have been disclosed in D'Hoore and Waibel is unreasonable.

Thus, the limitations of the present invention are not disclosed in the D'Hoore and Waibel and claims 6 and 15 are novel based on the features of D'Hoore and should be allowable.

Rejection of Claims 7 and 16

Referring to claims 7 and 16, the present invention teaches at least one uni-lingual anti-model generation engine, receiving a plurality of multi-lingual query commands to generate a plurality of uni-lingual anti-models corresponding to specific languages, and an anti-model combination engine, calculating the uni-lingual anti-models to generate the multi-lingual anti-models.

D'Hoore teaches using standard training techniques, training context independent models using Viterbi training of discrete density HMMs, constructing context dependent phoneme models, training the context dependent models using Viterbi training of discrete density HMMs, merging the context dependent and context independent phoneme models, and smoothing the trained context dependent models with the context independent models. Additionally, Waibel teaches applying the garbage models to model non-stationary noises.

Although D'Hoore and Waibel apply various models, the present invention applies the multi-lingual anti-models different from those model in D'Hoore and Waibel. Different models may result in different effects for speech recognition. Therefore, Applicant believes that the assertion by the Examiner that the models of the present invention have been disclosed in D'Hoore and Waibel to be unreasonable.

Further, the uni-lingual anti-model generation engine receiving the multi-lingual query commands to generate the plurality of uni-lingual anti-models and the anti-model combination engine calculating the uni-lingual anti-models to generate the multi-lingual anti-models are also not disclosed in D'Hoore and Waibel.

Thus, the limitations of the present invention are not disclosed in D'Hoore and Waibel and claims 7 and 16 are novel based on the features of D'Hoore and Waibel and should be allowable.

PCL/jcg

CONCLUSION

As presented in the above remarks, the present invention differs significantly from the references cited by the Examiner in the outstanding Office Action. None of these references, when taken alone or in combination, teaches all of the limitations recited in claims 1 and 9 of the present application. Therefore, Applicant believes that these claims are allowable over the cited references. Insofar as claims 2-7, depend from claim 1, and claims 10-16, depend from claim 9, these claims are similarly believed to be patentable over the cited references
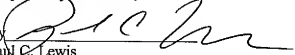
In view of the above Amendment, applicant believes the pending application is in condition for allowance.

Should there be any outstanding matters that need to be resolved in the present application, the Examiner is respectfully requested to contact Paul C. Lewis, Reg. No. 43,368 at the telephone number of the undersigned below, to conduct an interview in an effort to expedite prosecution in connection with the present application.

If necessary, the Commissioner is hereby authorized in this, concurrent, and future replies to charge payment or credit any overpayment to Deposit Account No. 02-2448 for any additional fees required under 37.C.F.R. §§1.16 or 1.147; particularly, extension of time fees.

Dated: May 12, 2008

Respectfully submitted,

By
Paul C. Lewis
Registration No.: 43,368
BIRCH, STEWART, KOLASCH & BIRCH, LLP
8110 Gatehouse Road
Suite 100 East
P.O. Box 747
Falls Church, Virginia 22040-0747
(703) 205-8000
Attorney for Applicant

PCL/jcg